

MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

RONALD FAGIN IBM Research Laboratory Sau Jose, California

GABRIEL M. KUPER<sup>1</sup> Stanford University Stanford, California

JEFFREY D. ULLMAN<sup>1</sup> Stanford University Stanford, California

MOSHE Y. VARDI<sup>2</sup> IBM Research Laboratory San Jose, California



Approved for modistribution

FILE CC

<sup>&</sup>lt;sup>1</sup> Work supported by AFOSR grant 80-0212

<sup>&</sup>lt;sup>2</sup>This work was partly done while this author was at Stanford University and supported by a Weizmann Fellowship and AFOSR grant 80 0212.

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE							
18. REPORT SECURITY CLASSIFICATION				1b. RESTRICTIVE MARKINGS			
UNCLASSIFIED							
28. SECURITY CLASSIFICATION AUTHORITY				3. DISTRIBUTION/AVAILABILITY OF REPORT			
26. DECLASSIFICATION/DOWNGRADING SCHEDULE				Approved for public release; distribution			
as. DECLASSIFICATION/DOWNGRADING SCHEDULE				unlimited.			
4. PERFORMING ORGANIZATION REPORT NUMBER(S)				5. MONITORING ORGANIZATION REPORT NUMBER(S)			
				AFOSR-TR. 16 1719			
6a. NAME OF PERFORMING ORGANIZATION			6b. OFFICE SYMBOL	7a. NAME OF MONITORING ORGANIZATION			
Chanfand University			(If applicable)	Air Force Office of Scientific Becomes			
Stanford University  6c. ADDRESS (City, State and ZIP Code)				Air Force Office of Scientific Research			
		· · · · · · · · · · · · · · · · · · ·	20	7b. ADDRESS (City, State and ZIP Code) Directorate of Mathematical & Information			
Department of Computer Science Stanford CA 94305				Sciences, Bldg 410, Bolling AFB DC 20332			
Stanz	CIU CA	94303		Scrences, br	ug 410, bo.	TITIE AL	B DC 20332
8a. NAME OF FUNDING/SPONSORING 8b. OFFICE SYMBOL (If applicable)				9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER			
				AFOSR-80-0212			
AFOSR NM  Sc. ADDRESS (City, State and ZIP Code)				10. SOURCE OF FUNDING NOS.			
				PROGRAM ELEMENT NO.	PROJECT NO.	TASK	WORK UNIT
Bldg 410, Bolling AFB DC 20332				61102F	2304	A7	
11. TITLE (Include Security Classification)				1			
UPDATING LOGICAL DATABASES					]		j
12. PERSONAL AUTHOR(S)							
Ronald Fagin*, Gabriel M. Kuper, Jeffrey Ullman, and Moshe Y. Vardi*.							
13a. TYPE OF REPORT 13b. TIME COVERED				14. DATE OF REPORT (Yr., Mo., Day) 15. PAGE COUNT 1984 13			
Technical FROM TO TO				1904			3
*IBM Research Laboratory, San Jose, California.							
17. COSATI CODES 18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)							um beri
FIELD	GROUP	SUB. GR.	1	, , , , , , , , , , , , , , , , , , , ,			
			]				
			<u> </u>	··· <del>-</del>			
19. ABSTRACT (Continue on reverse if necessary and identify by block number)  The authors suggest a new approach to database updates, in which a database is treated as a collection of theories. They investigate two issues: simultaneous multiple update operations, and equivalence of databases under update operations.							
20. DISTRIBUTION/AVAILABILITY OF ABSTRACT				21. ABSTRACT SECURITY CLASSIFICATION			
UNCLASSIFIED/UNLIMITED 🎞 SAME AS RPT. 🗆 DTIC USERS 🗖				UNCLASSIFIED			
22a. NAME OF RESPONSIBLE INDIVIDUAL				22b. TELEPHONE N (Include Area Co	ode)	22c OFFICE	SYMBOL
Dr. Robert N. Buchal				(3 <b>0</b> 2) <b>767–4</b> 93	39	NM	

DD FORM 1473, 83 APR

EDITION OF 1 JAN 73 IS OBSOLETE.

84 08 30 043

## Abstract

We suggest a new approach to database updates, in which a database is treated as a collection of theories. We investigate two issues: simultaneous multiple update operations, and equivalence of databases under update operations.

# 1. Introduction

One of the main problems in database theory is the problem of view updating, i.e., how to translate an update on a user view into an update of the database ([BS], [CA], [DB], [J], [Ke], [Kl], [O]). The problem is that in general there is no unique database update corresponding to the view update. Another problem is that of updating a database that must satisfy certain integrity constraints ([NY], [T]). The difficulty here is that the database after the update may no longer satisfy the constraints, in which case we may have to modify other things in the database, to ensure that the integrity constraints still hold. As in the case of view updates, there is not necessarily a unique way to modify the database so that the constraints still hold.

Fagin et al. [FUV] suggest that the appropriate framework for studying the semantics of updates is to treat the database as a consistent set of sentences in first-order logic, i.e., a theory. A theory is a description of the world, but is not necessarily a complete description; every model of the theory is a possible state of the world. Thus the database can be viewed as an exact description of our knowledge about the world. This framework was propounded in other papers (e.g., [Ko], [NG], [R]).

When one tries to update a theory by inserting or deleting some sentence, several new theories can accomplish the update. Fagin et al. [FUV] argue that we should try to minimize the change that is needed to accomplish the update. Unfortunately, even under this minimality constraint, there may be several theories that accomplish the update, with no reasonable way to choose between them. One approach to this, suggested in [FUV], is to define the result of the update to be the disjunction of all the possible theories that accomplish the update with minimal change. Two difficulties with this approach are that it requires us to have sentences of a rather complicated syntax, e.g., disjunctions of tuples in a relational database, and that the number of sentences in the database may grow doubly exponentially with each update.

The fact that several theories can accomplish a given update motivates an alternative approach: viewing the database as a collection of theories rather than a single theory. We call a collection of theories a flock. The advantage of this approach is that it is easier to deal with the multiplicity of flocks than with the multiplicity of theories. With the new approach, the sentences we get are of no greater complexity than those that were in the database or those that were inserted, and the number of sentences does not grow as fast as before.

In this paper, after presenting the two approaches to updates, databases as theories vs. databases as flocks, we investigate two basic issues. First we study batch operations, in which many sentences, rather than a single sentence, are inserted or deleted simultaneously. We then observe that two theories or flocks that are logically equivalent may not be equivalent after an update is performed. We give necessary and sufficient conditions for equivalence forever, i.e., equivalence that is preserved under updates.

ATR PORCE OFFICE OF COLUMN TO A COLUMN TO

# 2. Updates of Theories.

Our basic units of information are sentences, i.e. formulas without free variables. We do not allow inconsistent sentences, and we do not allow the deletion of valid sentences. A theory is a consistent set of sentences. We shall use the letters S and T to denote theories, and the letters  $\sigma$  and  $\tau$  to denote sentences.

We start by describing the framework developed in [FUV].

#### Definition 1:

- 1. A theory T accomplishes the deletion of  $\sigma$  from S if  $T \not\models \sigma$ .
- 2. A theory T accomplishes the insertion of  $\sigma$  into S if  $\sigma \in T$ .

**Definition 2:** Let  $T_1$ ,  $T_2$  and T be theories.

- 1.  $T_1$  has fewer insertions than  $T_2$  with respect to T if  $T_1 T \subset T_2 T$ .
- 2.  $T_1$  has fewer deletions than  $T_2$  with respect to T if  $T T_1 \subset T T_2$ .
- 3.  $T_1$  has fewer changes than  $T_2$ , with respect to T, if  $T_1$  has fewer deletions than  $T_2$ , or  $T_1$  and  $T_2$  have the same deletions  $(T T_1 = T T_2)$  and  $T_1$  has fewer insertions than  $T_2$ .

**Definition 3:** A theory T accomplishes an update u of S minimally if T accomplishes u and there is no theory T' that accomplishes u and has fewer changes than T with respect to S.

Theorem 1: ([FUV]) Let S and T be theories and let  $\sigma$  be a sentence. Then

- 1. S accomplishes the deletion of  $\sigma$  from T minimally iff S is a maximal subset of T that is consistent with  $\neg \sigma$ .
- 2.  $S \cup \{\sigma\}$  accomplishes the insertion of  $\sigma$  into T minimally iff S is a maximal subset of T that is consistent with  $\sigma$ .

There could be many theories that accomplish an update minimally. Suppose that  $T_1, \ldots, T_n$  are the theories that accomplish an update u of S minimally. It is argued in [FUV] that the result of u should be a theory T such that

$$\operatorname{Mod}(T) = \bigcup_{1 \leq i \leq n} \operatorname{Mod}(T_i),$$

where Mod(S) is the set of models of the theory S.

**Definition 4:** Let  $T_1, \ldots, T_n$  be theories. The disjunction of these theories is defined to be the theory

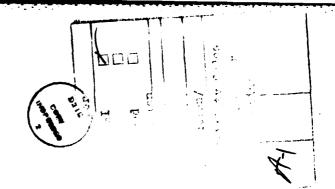
$$\bigvee_{1 \le i \le n} T_i = \{ \tau_1 \vee \cdots \vee \tau_n \mid \tau_i \in T_i, 1 \le i \le n \}.$$

It is shown in [FUV] that

$$\operatorname{Mod}\left(\bigvee_{1\leq i\leq n}T_i\right)=\bigcup_{1\leq i\leq n}\operatorname{Mod}(T_i).$$

Thus they suggest that if  $T_1, \ldots, T_n$  are the theories that accomplish an update u minimally, then the result of u should be  $\bigvee_{1 \le i \le n} T_i$ .

<sup>&</sup>lt;sup>3</sup>We use  $\leq$  to denote inclusion, and  $\in$  to denote proper inclusion.



## 3. Flocks.

In this section, we shall describe another approach to updates, namely using collections of theories. We call these collections flocks. The intuitive idea is that since we have many possible theories that accomplish an update minimally, we reflect this ambiguity by keeping all these theories.

Definition 5: A flock S is a set of theories. The models of S are

$$\operatorname{Mod}(S) = \bigcup_{S \in \mathbf{S}} \operatorname{Mod}(S).$$

To update a flock we have to update each theory in the flock. Formally:

**Definition 6:** Let  $S = \{S_1, \ldots, S_n\}$  be a flock. A flock  $T = \{T_1, \ldots, T_n\}$  accomplishes an update u of S minimally if  $T_i$  accomplishes the update of  $S_i$  minimally, for  $1 \le i \le n$ .

Again, there could be many flocks that accomplish an update minimally. Suppose that  $T_1, \ldots, T_n$  are the flocks that accomplish an update u of S minimally. As in [FUV], we contend that the result of u should be a flock T such that

$$\operatorname{Mod}(\mathbf{T}) = \bigcup_{1 \leq i \leq n} \operatorname{Mod}(\mathbf{T}_i).$$

It is easy to show that the flock  $\bigcup_{1 \le i \le n} \mathbf{T}_i$  has this property. This motivates the following definition:

**Definition 7:** Let S be a flock, and let  $S_1, \ldots, S_n$  be the flocks that accomplish an update u of S minimally. Then the result of u is the flock  $\bigcup_{1 \le i \le n} S_i$ .

**Lemma 2:** Let  $S = \{S_1, \ldots, S_n\}$  be a flock. For each theory  $S_i$ , let  $S_i^1, \ldots, S_i^{j_i}$  be the theories that accomplish the update u of  $S_i$  minimally. Then the result of applying u to S is the flock

$$\mathbf{S}' = \{S_i^k \mid 1 \leq i \leq n, 1 \leq k \leq j_i\}.$$

**Proof:** Let S' be the result of the update. If  $S \in S'$  then, by Definition 7,  $S \in S_j$ , for some  $S_j$  that accomplishes the update minimally. But then, by Definition 6, S accomplishes the update of some  $S_i \in S$  minimally, i.e., S is one of the theories  $S_i^1, \ldots, S_i^{j_i}$ .

Now let  $S = S_i^k$  for some  $k, 1 \le k \le j_i$ . Then S accomplishes the update u of  $S_i$  minimally. For each j,  $1 \le j \le n, j \ne i$ , let  $S^j$  be any theory that accomplishes the update u of  $S_j$  minimally. Then, by Definition 6, the flock  $\{S^1, \ldots, S^{j-1}, S, S^{j+1}, \ldots, S^n\}$  accomplishes the update u of S minimally and so, by Definition 7, each theory in this flock is in S'. In particular,  $S \in S'$ .

In other words, to update a flock, consider each theory in the flock in turn. Take all theories that accomplish the update minimally and put them in the new flock.

Note that if a flock is a singleton, i.e., contains exactly one theory, its models as a theory and as a flock are the same. Also, the flock that we get after applying an update to such a flock has the same models as the theory we get by applying the update to the single member of that flock, as the following lemma shows.

**Lemma 3:** Let  $S = \{S\}$  be a singleton flock, and u an update. If S' is the result of applying the update u to the theory S, and S' is the result of applying u to the flock S, then S' and S' have the same models.

**Proof:** By Lemma 2, the result of applying u to S is the flock  $S' = \{S_i \mid 1 \le i \le j\}$ , where  $S_1, \ldots, S_j$  are the theories that accomplish the update u of S minimally. Similarly, the result of applying u to S is the

theory  $S' = \bigvee_{1 \le i \le j} S_i$ . By Definition 5 and the comments at the end of Section 2, S' and S' have the same models.

Even though the result of an update has the same models under both approaches, under future updates their results may differ, as the following example shows.

**Example 1:** If we start with the flock  $\{\{A,B\}\}$ , and delete  $A \wedge B$  from it using the flocks approach, we take all the maximal subtheories of  $\{A,B\}$  that do not imply  $A \wedge B$ , namely  $\{A\}$  and  $\{B\}$ . That is, the resulting flock is  $\{\{A\},\{B\}\}\}$ . If we now delete A and then delete B, we end up with the flock containing only the empty theory, i.e., anything is a model of the result. On the other hand, if we start with the theory  $\{A,B\}$ , and delete  $A \wedge B$ , we get the theory  $\{A \vee B\}$ . This has the same models as the flock  $\{\{A\},\{B\}\}\}$ . However if we now delete A and then delete B, we still have the theory  $\{A \vee B\}$ , which does not have the same models as the empty theory.

In practice, singleton flocks are the most likely to be used as the starting state of the database (in fact, the starting state will probably be  $\{\emptyset\}$ ). It would be interesting to characterize the flocks that are obtained from singleton flocks by a sequence of update operations. Another interesting question is the comparative merit of the two approaches: theories vs. flocks. We know that these approaches yield different results for the same updates. Which one of them is more correct?

# 4. Batch Operations.

Batch operations consist of deleting or inserting several sentences simultaneously.

**Definition 8:** Let S be a theory and let  $\Sigma$  be a set of sentences. We say that S' accomplishes the deletion of  $\Sigma$  from S if S'  $\not\models \sigma$  for each  $\sigma \in \Sigma$ . We say that S' accomplishes the insertion of  $\Sigma$  into S if  $\Sigma \subseteq S'$ . We say that S' accomplishes an update u of S minimally if S' accomplishes u and there is no theory that accomplishes u with fewer changes.

The above definition is non-constructive in the sense that it does not explicitly say how to find those theories that accomplish an update minimally. The following theorem gives a constructive equivalent condition, which generalizes Theorem 2.

**Theorem 4:** Let S and T be theories and  $\Sigma$  a set of sentences. Then

- 1. S accomplishes the deletion of  $\Sigma$  from T minimally iff S is a maximal subset of T such that  $S \cup \{\neg \sigma\}$  is consistent for all  $\sigma$  in  $\Sigma$ .
- 2.  $S \cup \Sigma$  accomplishes the insertion of  $\Sigma$  into T minimally iff S is a maximal subset of T that is consistent with  $\Sigma$ .

#### Proof:

- 1. If S is a maximal subset of T that is consistent with  $\neg \sigma$  for every  $\sigma \in \Sigma$ , then clearly S accomplishes the deletion of  $\Sigma$  from S. Assume that S does not accomplish the deletion minimally, i.e., there is a theory S' that accomplishes the deletion with fewer changes than S with respect to T. If S' has fewer deletions than S, then  $T S' \subset T S$ . But then  $S' \cap T$  is also consistent with  $\neg \sigma$ , for all  $\sigma$  in  $\Sigma$ , contrary to the maximality of S. Therefore S' must have the same deletions as S with respect to T. Clearly S' cannot have fewer insertions than S, since S has no such insertions at all.
  - If S accomplishes the deletion minimally, it must be consistent with  $\neg \sigma$  for every  $\sigma \in \Sigma$ . It is also clear that  $S \subseteq T$  since if it contained sentences not in T we could remove them and get a theory that accomplished the update with the same deletions and with fewer insertions. If S is not a maximal subset of T that is consistent with all the  $\neg \sigma$ 's, then there is a theory that accomplishes the update with fewer deletions than S.
- 2. Let S be a maximal subset of T that is consistent with  $\Sigma$ .  $S \cup \Sigma$  clearly accomplishes the insertion of  $\Sigma$ . Suppose that S' accomplishes the update with fewer deletions, and let  $S'' = S' \cap T$ . Then  $T S'' = T S' \subset T S$ , and therefore  $S \subset S'' \subseteq T$  and S'' is consistent with  $\Sigma$ -contradicting the maximality of S. Clearly, no theory can accomplish the insertion with the same deletions and with fewer insertions than  $S \cup \Sigma$ , since the only insertions are  $\Sigma$ .
  - If  $S \cup \Sigma$  accomplishes the insertion of  $\Sigma$  minimally, we must have  $S \subseteq T$  and S consistent with  $\Sigma$ . If S is not a maximal subset of T that was consistent with  $\Sigma$ , then we can find S' consistent with  $\Sigma$  that satisfies  $S \subset S' \subseteq T$ . But then  $S' \cup \Sigma$  accomplishes the insertion with fewer deletions—contradiction.

Using Definition 8, we can define the result of batch updates both for theories and for flocks. For theories, we define the result of the update to be the disjunction of all the theories that accomplish the update minimally, as in Definition 4. For flocks, we use Definitions 6 and 7. Namely, to update a flock consider each theory in the flock in turn, take all theories that accomplish the update of this theory minimally, and put them into the new flock. In the sequel, we reserve the term update (respectively, deletion, insertion) for the case where a single sentence is deleted or inserted, to distinguish it from batch update (respectively, batch deletion, batch insertion), where a set of sentences is deleted or inserted.

The following example shows that the batch deletion of  $\Sigma$  does not always give the same result as deleting the sentences in  $\Sigma$  one by one.

**Example 2:** Deleting  $\{A, B\}$  from the theory  $\{A, B, A \equiv B\}$  results in the theory  $\{A \equiv B\}$ . If, on the other hand, we delete first A, we get the theory  $\{B \vee (A \equiv B)\}$ , which remains unchanged after deleting B. Deleting first B and then A gives us the theory  $\{A \vee (A \equiv B)\}$ .

Deleting  $\{A, B\}$  from the flock  $\{\{A, B, A \equiv B\}\}$  results in the flock  $\{\{A \equiv B\}\}$ . If, on the other hand, we delete first A we get the flock  $\{\{B\}, \{A \equiv B\}\}$ , and if we then delete B we end up with the flock  $\{\emptyset, \{A \equiv B\}\}$ . This is different from the flock  $\{\{A \equiv B\}\}$ , since the union of the models of the first flock consists of all possible structures, whereas the models of  $\{\{A \equiv B\}\}$  are only those models in which A and B are equivalent.

Similarly, the insertion of  $\Sigma$  does not give the same result as inserting the sentences in  $\Sigma$  one by one. The following theorem shows, however, that for flocks, batch insertions can be simulated by single updates.

**Theorem 5:** Let  $\Sigma = \{\sigma_1, \dots, \sigma_n\}$  be a consistent set of sentences and let **S** be a flock. Then the result of inserting  $\Sigma$  into **S** is the same as first deleting  $\neg(\sigma_1 \wedge \dots \wedge \sigma_n)$  and then inserting the  $\sigma_i$ 's one by one.

**Proof:** A theory S is consistent with  $\Sigma$  iff  $S \not\models \neg(\sigma_1 \wedge \cdots \wedge \sigma_n)$ . Let **T** be the result of deleting  $\neg(\sigma_1 \wedge \cdots \wedge \sigma_n)$  from S. We claim that  $S' = \{T \cup \Sigma \mid T \in T\}$  is the result of inserting  $\Sigma$  into S.

First, let T be a theory in T. It accomplishes the deletion of  $\neg(\sigma_1 \land \cdots \land \sigma_n)$  from some  $S \in S$  minimally. We claim that  $T \cup \Sigma$  accomplishes the insertion of  $\Sigma$  into S minimally. It is clear that  $T \cup \Sigma$  accomplishes the insertion of  $\Sigma$ . If T' accomplishes the insertion with fewer deletions that  $T \cup \Sigma$ , then T' also accomplishes the deletion of  $\neg(\sigma_1 \land \cdots \land \sigma_n)$  from S with fewer deletions than T with respect to S, a contradiction. Clearly no theory can accomplish the insertion of  $\Sigma$  into S with fewer insertions that  $T \cup \Sigma$  with respect to S, since the only insertions here are the sentences of  $\Sigma$ . This shows that each theory in S' is in the result of inserting  $\Sigma$  into S.

Now let T be a theory in the result of inserting  $\Sigma$  into S, i.e., T accomplishes the insertion of  $\Sigma$  into some  $S \in S$  minimally. Let  $T' = T - (\Sigma - S)$ . Then T' is consistent with  $\Sigma$  and so T' accomplishes the deletion of  $\neg(\sigma_1 \wedge \cdots \wedge \sigma_n)$  from S. If some theory S' accomplishes the deletion with fewer deletions than T' with respect to S, then  $S' \cup \Sigma$  accomplishes the insertion of  $\Sigma$  with fewer deletions than T, a contradiction. Therefore, T' accomplishes the deletion of  $\neg(\sigma_1 \wedge \cdots \wedge \sigma_n)$  from S minimally, i.e.,  $T' \in T$ , and therefore  $T = T' \cup \Sigma$  is in S'.

## Remarks:

1. The theorem does not hold for theories. For example, let S be the theory  $\{A, B\}$ , and  $\Sigma$  the set  $\{(A \not\equiv B) \land C\}$ . Then the result of inserting  $\Sigma$  into S is the theory

$$\{A \lor ((A \neq B) \land C), B \lor ((A \neq B) \land C), A \lor B, (A \neq B) \land C\}.$$

On the other hand, the result of deleting  $(A = B) \vee \neg C$  from S is the theory  $\{A \vee B\}$ , and if we then just insert  $(A \neq B) \wedge C$ , we get the theory  $\{A \vee B, (A \neq B) \wedge C\}$ .

2. There are batch deletions from flocks that cannot be simulated by any sequence of single updates. For example, if we delete  $\{A, B\}$  from the flock  $\{\{A \vee B, A \vee \neg B, \neg A \vee B\}\}$ , we get the flock  $\{\{A \vee B\}, \{A \vee \neg B, \neg A \vee B\}\}$ . It is shown in [Ku] that the latter flock cannot be obtained from any singleton flock by single updates.

# 5. Equivalence Forever.

## 5.1. Definitions.

Two theories or flocks are logically equivalent if they have the same models. Nevertheless, this does not guarantee that they will continue to have the same models after any sequence of updates, as the next example shows.

**Example 3:** The two theories  $\{B\}$  and  $\{B, A \vee B\}$  are logically equivalent. However, if we delete B from both of them we get the nonequivalent theories  $\emptyset$  and  $\{A \vee B\}$ .

The two flocks  $\{\{B\}\}\$  and  $\{\{B, A \lor B\}\}\$  are logically equivalent. After deleting B from both of them we get the nonequivalent flocks  $\{\emptyset\}\$  and  $\{\{A \lor B\}\}\$ .

We say that two theories or flocks are equivalent forever if after applying any sequence of updates we always get two theories or flocks that have the same models. In the rest of this section we supply characterizations for equivalence forever.

We use the following definition.

**Definition 9:** We say that a theory S covers a theory T iff every sentence  $\tau$  in T is logically equivalent to a conjunction  $\sigma_1 \wedge \cdots \wedge \sigma_n$  of sentences in S. (An empty conjunction is by convention valid.)

## 5.2. Equivalence Forever for Theories.

Theorem 6: Let S and T be finite theories. The following are equivalent.

- 1. S and T are equivalent forever under updates.
- 2. S and T are equivalent forever under batch updates.
- 3. S and T are equivalent forever under deletions.
- 4. S and T are equivalent forever under batch deletions.
- 5. Each subset of S is logically equivalent to a subset of T, and vice-versa.
- 6. S covers T, and vice-versa.

**Proof:** (2)  $\Rightarrow$  (1), (2)  $\Rightarrow$  (4), (4)  $\Rightarrow$  (3), (5)  $\Rightarrow$  (6) and (1)  $\Rightarrow$  (3) are obvious. We shall show (3)  $\Rightarrow$  (6), (6)  $\Rightarrow$  (5) and (5)  $\Rightarrow$  (2).

- (3)  $\Rightarrow$  (6) We shall prove the following statement, which we call statement (\*), inductively on k:
  - (\*) Let  $T_1$  and  $T_2$  be finite theories that are equivalent forever under deletions. If there is a structure M that obeys  $\tau \in T_1$  and that also obeys exactly k sentences in  $T_2$ , then  $\tau$  is equivalent to a conjunction of sentences in  $T_2$ .

Statement (\*) implies (3)  $\Rightarrow$  (6). For, let  $S = T_1$  and  $T = T_2$ , and let  $\tau$  be an arbitrary member of S. Let M be a structure which obeys  $\tau$ . (There is such a structure since we deal only with consistent sentences.) Since T is finite, there is some k (possibly k = 0) such that M obeys exactly k sentences in T. Then statement (\*) tells us that  $\tau$  is equivalent to a conjunction of sentences in T, as desired.

If  $T_1$  and  $T_2$  are finite theories and if M is a structure, it is convenient for us to define  $\sigma(M, T_1, T_2)$  to be the sentence

$$\bigvee \{\sigma \mid \sigma \in T_1 \cup T_2 \text{ and } M \text{ violates } \sigma \}.$$

It is easy to see that there is a single maximal theory which results from deleting this sentence from  $T_1$ , namely, the set of all sentences in  $T_1$  which are true in M. Of course, the same is true about  $T_2$ . We are now ready to prove statement (\*), by induction on k.

k=0: In this case, M obeys no sentence in  $T_2$ . Let us denote by  $T_1'$  (respectively,  $T_2'$ ) the result of deleting  $\sigma(M,T_1,T_2)$  from  $T_1$  (respectively,  $T_2$ ). Since M obeys no sentence in  $T_2$ , it follows that  $T_2'$  is the empty theory, which every structure obeys. Since  $T_1'$  and  $T_2'$  are equivalent (by equivalence forever of  $T_1$  and  $T_2$  under deletions), it follows that  $T_1'$  consists of valid sentences. But  $\tau$  belongs to  $T_1'$ , since M obeys  $\tau$ . It follows that  $\tau$  is valid, and is therefore equivalent to a conjunction of sentences in  $T_2$ .

Inductive step: Assume that the inductive hypothesis (\*) holds, with k' substituted for k, for every k' < k, and for every choice of  $T_1$  and  $T_2$ . Let  $T_1$  and  $T_2$  be finite theories that are equivalent forever under deletions, and let M be a structure which obeys  $\tau \in T_1$  and which also obeys exactly k sentences in  $T_2$ . We must show that  $\tau$  is equivalent to a conjunction of sentences in  $T_2$ .

Let us denote by  $T_1'$  (respectively,  $T_2'$ ) the result of deleting  $\sigma(M, T_1, T_2)$  from  $T_1$  (respectively,  $T_2$ ). Then  $T_1'$  is a subset of  $T_1$  which contains  $\tau$ , and  $T_2'$  is a subset of  $T_2$  which contains exactly k sentences (namely, those sentences in  $T_2$  which are true in M). By equivalence forever of  $T_1$  and  $T_2$  under deletions, we know that  $T_1'$  and  $T_2'$  are also equivalent forever under deletions. In particular,  $T_1'$  and  $T_2'$  are equivalent, and so  $T_2'$  implies  $\tau$ . If also  $\tau$  were to imply  $T_2'$  (that is, if  $\tau$  were to imply every member of  $T_2'$ ), then we would be done, since  $\tau$  would be equivalent to the subset  $T_2'$  of  $T_2$ . So we can assume that  $\tau$  does not imply  $T_2'$ . Therefore, there is a structure M' which obeys  $\tau$  but not  $T_2'$ . Let k' be the number of members of  $T_2'$  which M' obeys. Then  $0 \le k' < k$ , since  $T_2'$  contains k sentences, not all of which M' obeys. By inductive hypothesis (\*), where  $T_1'$ ,  $T_2'$ , and k' play the roles of  $T_1$ ,  $T_2$ , and k respectively, it follows that  $\tau$  is equivalent to a conjunction of members of  $T_2'$ , and hence of  $T_2$ .

- (6) ⇒ (5) Let S' be a subset of S. For each σ ∈ S', let T<sub>σ</sub> be a subset of T such that σ is equivalent to the conjunction of members of T<sub>σ</sub>. Let T' be the union of all sets T<sub>σ</sub> where σ ∈ S'. We now show that S' is equivalent to T'. If τ ∈ T', find σ ∈ S' such that τ ∈ T<sub>σ</sub>. Then σ implies τ, so S' implies τ. Hence, S' implies T'. Conversely, assume that σ ∈ S'. Then T<sub>σ</sub> implies σ, and so T' implies σ. Hence, T' implies S'.
- (5)  $\Rightarrow$  (2) Assume that (5) holds. It suffices to show that if  $S^{(1)}$  (respectively,  $T^{(1)}$ ) is the result of applying a batch update u to S (respectively, T), then every subset of  $S^{(1)}$  is equivalent to a subset of  $T^{(1)}$  and vice versa.

Let the update u be the deletion of  $\Sigma$  (we shall remark at the end how to modify the proof to deal with the case where u is an insertion.) Let  $S^*$  be a maximal subset of S that is consistent with  $\neg \sigma$ , for every  $\sigma \in \Sigma$ . By assumption, there is a subset of T that is equivalent to  $S^*$ . Let  $T^*$  be a maximal such subset of T. We now show that  $T^*$  is a maximal subset of T that is consistent with  $\neg \sigma$ , for every  $\sigma \in \Sigma$ . Clearly  $T^*$  is consistent with every  $\neg \sigma$ , since  $T^*$  is equivalent to  $S^*$ , and  $S^{\dagger}$  is consistent with every  $\neg \sigma$ . If  $T^{\dagger}$  is not maximal, then find  $T' \subseteq T$  consistent with every  $\neg \sigma$  such that  $T^* \subset T'$ . By definition of  $T^{\dagger}$ , we know that  $T^*$  is not equivalent to T'. By hypothesis, there is a subset of S that is equivalent to T'; let S' be a maximal such subset. Since  $S^{\dagger} \equiv T^{\dagger} \subset T' \equiv S'$ , it follows that S' implies  $S^{\dagger}$ , and so  $S' \cup S^{\dagger}$  is equivalent to S'. By maximality of S', it follows that  $S' \subseteq S'$ . But  $S^{\dagger} \neq S'$ , since  $S^{\dagger} \equiv T^{\dagger} \neq T' \equiv S'$ . Hence,  $S^{\dagger} \subset S'$ . Since T' is consistent with  $\neg \sigma$ , for every  $\sigma \in \Sigma$ , and since  $T' \equiv S'$ , it follows that S' is consistent with every  $\neg \sigma$ . This contradicts maximality of  $S^{\dagger}$ .

Let us call each maximal subset of S (respectively, T) that is consistent with every  $\neg \sigma$  an S-candidate (respectively, a T-candidate). We have shown that for each S-candidate S<sup>†</sup> there is a

T-candidate  $T^{\dagger}$  such that  $S^{\dagger} \equiv T^{\dagger}$ . Similarly, for each T-candidate  $T^{\dagger}$  there is an S-candidate  $S^{\dagger}$  such that  $S^{\dagger} \equiv T^{*}$ . Furthermore, it is easy to show that this correspondence is bijective. That is, if  $S_{1}, \ldots, S_{n}$  are all of the distinct S-candidates, then there is a listing  $T_{1}, \ldots, T_{n}$  of all of the distinct T-candidates such that  $S_{i} \equiv T_{i}$ , for  $1 \leq i \leq n$ .

The result  $S^{(1)}$  of the update on S is the theory  $\bigvee \{S_i \mid 1 \leq i \leq n\}$ , and analogously for  $T^{(1)}$ . We shall show  $T^{(1)}$  covers  $S^{(1)}$ . As in the proof that  $(4) \Rightarrow (3)$ , it then follows that every subset of  $S^{(1)}$  is equivalent to a subset of  $T^{(1)}$ , as desired.

Let  $\alpha$  be a member of  $S^{(1)}$ . We know that  $\alpha$  is of the form  $\alpha_1 \vee \cdots \vee \alpha_n$ , where  $\alpha_i \in S_i$ , for  $1 \leq i \leq n$ . By assumption, there is a subset  $T_i'$  of T which is equivalent to  $\alpha_i$ . Since  $\alpha_i \in S_i \equiv T_i$ , it follows that  $T_i$  implies  $\alpha_i$ , and hence  $T_i$  implies  $T_i'$ . So by maximality of  $T_i$ , we know that  $T_i' \subseteq T_i$ . Let Q be the set  $\{\tau_1 \vee \cdots \vee \tau_n \mid \tau_i \in T_i' \text{ for } 1 \leq i \leq n\}$ . Then  $Q \subseteq T^{(1)}$ . Let  $\tau$  be the conjunction of members of Q. The proof is complete if we show that  $\alpha$  is equivalent to  $\tau$ . Let  $\tau_i'$  be the conjunction of members of  $T_i'$ , for  $1 \leq i \leq n$ , and let  $\gamma$  be the disjunction  $\tau_1' \vee \cdots \vee \tau_n'$ . Clearly  $\tau_i'$  is equivalent to  $\alpha_i$ , since both are equivalent to  $T_i'$  ( $1 \leq i \leq n$ ). Hence,  $\gamma$  is equivalent to  $\alpha$ . But  $\tau$  is the conjunctive normal form of  $\gamma$ , and consequently  $\tau$  is equivalent to  $\alpha$ .

We close by remarking how the proof should be modified to deal with insertions rather than deletions. Assume that the update u is the insertion of  $\Sigma$ . Let us call each maximal subset of S (respectively, T) that is consistent with  $\Sigma$  an S-candidate (respectively, a T-candidate). Just as before, it follows that if  $S_1, \ldots, S_n$  are all of the distinct S-candidates, then there is a listing  $T_1, \ldots, T_n$  of all of the distinct T-candidates such that  $S_i \equiv T_i$ , for  $1 \le i \le n$ . The result  $S^{(1)}$  of the update on S is the theory  $\bigvee \{S_i \cup \Sigma \mid 1 \le i \le n\}$ , and analogously for  $T^{(1)}$ . Let  $\alpha$  be a member of  $S^{(1)}$ . We know that  $\alpha$  is of the form  $\alpha_1 \vee \cdots \vee \alpha_n$ , where  $\alpha_i \in S_i \cup \Sigma$ , for  $1 \le i \le n$ . If no  $\alpha_i$  is in  $\Sigma$ , then the proof proceeds as before. Assume now that some  $\alpha_i$  is in  $\Sigma$ . For simplicity in description, assume that  $\alpha_1$  is in  $\Sigma$ , but  $\alpha_i \in S_i$  for  $1 \le i \le n$  (otherwise there is an obvious modification in the proof). As before, find a subset  $T_i$  of T which is equivalent to T and T is equivalent to the conjunction of members of T.

## 5.3. Equivalence Forever of Flocks.

We do not have, at present, a simple necessary and sufficient condition for equivalence forever of general flocks. However, for singleton flocks, i.e., flocks that contain only one theory, we can prove an analogue to Theorem 6.

**Theorem 7:** Let S and T be finite theories, and let  $S = \{S\}$  and  $T = \{T\}$  be singleton flocks. The following are equivalent.

- 1. S and T are equivalent forever under updates.
- 2. S and T are equivalent forever under batch updates.
- 3. S and T are equivalent forever under deletions.
- 4. S and T are equivalent forever under batch deletions.
- 5. Each subset of S is logically equivalent to a subset of T, and vice-versa.
- 6. S covers T and T covers S.

**Proof:** (2)  $\Rightarrow$  (1), (2)  $\Rightarrow$  (4), (4)  $\Rightarrow$  (3), (5)  $\Rightarrow$  (6) and (1)  $\Rightarrow$  (3) are obvious, and (6)  $\Rightarrow$  (5) was proven in Theorem 6. We now show (3)  $\Rightarrow$  (6) and (6)  $\Rightarrow$  (2).

(3)  $\Rightarrow$  (6) Assume that S does not cover T. Then there is a sentence  $\tau$  in T that is not logically equivalent to any conjunction of sentences of S. Let  $\Sigma$  be the set of sentences in  $S \cup T$  that are not implied by  $\tau$ . Let R be the set of maximal disjunctions of sentences in  $\Sigma$ , i.e., the set R of all disjunctions of sentences in  $\Sigma$  such that if we add any other sentence in  $\Sigma$  to the disjunction, the result is implied by  $\tau$ . Formally, R consists of all sentences of the form  $\sigma_1 \vee \cdots \vee \sigma_k$ , where each  $\sigma_i$  is in  $\Sigma$ ,

$$\tau \not\models \sigma_1 \vee \cdots \vee \sigma_k$$

and if  $\sigma$  is any sentence in  $\Sigma$  distinct from all the  $\sigma_i$ 's, then

$$\tau \models \sigma_1 \lor \cdots \lor \sigma_k \lor \sigma.$$

We now show that if we delete the sentences in R from the flock  $S = \{S\}$ , one by one, in any order, the resulting flock S' will be equal to  $\{S - \Sigma\}$ , and similarly deleting R from  $T = \{T\}$  will result in  $\{T - \Sigma\}$ . We prove this for S, and an analogous proof holds for T.

Since no sentence in  $\Sigma$  is implied by  $\tau$ , every  $\sigma$  in  $\Sigma$  can be extended to a maximal disjunction  $\sigma \vee \sigma_1 \vee \cdots \vee \sigma_k$  that is in the set R. After deleting this disjunction, we get a flock of theories, none of which can contain any of the sentences  $\sigma$ ,  $\sigma_1$ , ...,  $\sigma_k$ . Therefore after deleting all of the sentences in R from S we get a flock S' of theories, each of which must be a subset of  $S - \Sigma$ .

We now show by induction on the number of deletions that the result is a singletor flock, consisting of one theory that is a superset of  $S - \Sigma$ . The basis for the induction is the initial flock S. We now show that if we have a flock consisting of one theory that is a superset of  $S - \Sigma$  and a subset of S and we delete a sentence in the set R from it, we get a singleton flock that also consists of one theory that is a superset of  $S - \Sigma$  and a subset of S.

Suppose that we have such a flock consisting of the theory S' and we delete from it a sentence  $\sigma_1 \vee \cdots \vee \sigma_k$  of R. Let  $\Sigma = \{\sigma_1, \ldots, \sigma_k\} = \{\alpha_1, \ldots, \alpha_m\}$ . Since  $\sigma_1 \vee \cdots \vee \sigma_k$  is maximal, r implies  $\sigma_1 \vee \cdots \vee \sigma_k \vee \alpha_i$  for  $1 \leq i \leq m$ , and consequently  $\neg \alpha_i$  implies  $\sigma_1 \vee \cdots \vee \sigma_k \vee \neg \tau$  for  $1 \leq i \leq m$ . Suppose that  $\{\tau, \alpha_1, \ldots, \alpha_m\}$  implies  $\sigma_1 \vee \cdots \vee \sigma_k$ , then  $\tau$  implies  $\sigma_1 \vee \cdots \vee \sigma_m \vee \sigma_1 \vee \cdots \vee \sigma_k$ . But then it follows that  $\tau$  implies  $\sigma_1 \vee \cdots \vee \sigma_k$  - contradiction. But  $\tau$  implies  $S - \Sigma$  by definition and S' is a subset of S, so  $\{S' - \Sigma\} \cup \{\alpha_1, \ldots, \alpha_m\}$  does not imply  $\sigma_1 \vee \cdots \vee \sigma_k$ . Thus  $S' = \{\sigma_1, \ldots, \sigma_k\}$  is consistent with  $\neg(\sigma_1 \vee \cdots \vee \sigma_k)$ . It follows that the result of deleting  $\sigma_1 \vee \cdots \vee \sigma_k$  from  $\{S'\}$  is  $\{S' = \{\sigma_1, \ldots, \sigma_k\}\}$ . This completes the induction and shows that the result of deleting R from  $S = \{S\}$  is  $S' = S - \Sigma$ .

By the definition of  $\Sigma$ , we have  $\tau \models S - \Sigma$ , and therefore  $\tau$  implies the conjunction of all the sentences in  $S - \Sigma$ . Since  $\tau$  is not logically equivalent to any conjunction of a collection of sentences in S, it follows that  $S - \Sigma \not\models \tau$ . Therefore, there must be a model M of  $S - \Sigma$  that is not a model of  $\tau$ . Then M is a model of S'. However, since  $\tau$  in in  $T - \Sigma$ , M is not a model of T'. It follows that S and T are not equivalent forever under deletions.

(6)  $\Rightarrow$  (2) We show by induction on the number of updates that we always have

$$(\forall S' \in \mathbf{S}')(\exists T' \in \mathbf{T}')(S' \text{ covers } T' \land T' \text{ covers } S'), \tag{1}$$

where S' and T' are the flocks we get from S and T by performing some updates. By our assumption, Condition 1 holds at the beginning, when both flocks are singletons.

Assume that Condition 1 holds after some insertions and deletions. We have to show that it continues to hold after deleting a set of sentences  $\Sigma$ . The argument for insertion is similar and is

left to the reader. We shall use  $S^1$  and  $T^1$  for the flocks before the deletion,  $S^2$  and  $T^2$  for the flocks afterwards.

Let  $S^2$  be a theory in the flock  $S^2$ . We first show that there is some theory in  $T^2$  that covers  $S^2$ . By the definition of deletion,  $S^2$  must be a maximal subset of some theory  $S^1$  in the flock  $S^1$  that does not imply any sentence in  $\Sigma$ . By the inductive hypothesis, there is a theory  $T^1$  in the flock  $T^1$  such that  $S^1$  covers  $T^1$ , and  $T^1$  covers  $S^1$ . Let  $\sigma_i$  be any sentence in the theory  $S^2$ . Since  $S^2$  is a subset of  $S^1$  and  $T^1$  covers  $S^1$ , there are sentences  $\tau_{i1}, \ldots, \tau_{im_i}$  in  $T^1$  such that  $\sigma_i \equiv \tau_{i1} \wedge \cdots \wedge \tau_{im_i}$ .

Let A be the set of all these  $\tau_{ij}$ 's, for all  $\sigma_i$ 's in  $S^2$ . We claim that A does not imply any sentence in  $\Sigma$ . Assume otherwise, i.e.,  $A \models \sigma$ , for some  $\sigma$  in  $\Sigma$ . Since each  $\sigma_i$  in  $S^2$  implies all the corresponding  $\tau_{ij}$ 's in A, we have  $S^2 \models A$ , and therefore  $S^2 \models \sigma$ , a contradiction. Therefore A does not imply any sentence in  $\Sigma$  and can be extended to a maximal subset of  $T^1$  with this property. Call the maximal subset  $T^2$ . Since A covers  $S^2$ ,  $T^2$  also covers  $S^2$ . We shall now show that  $S^2$  covers  $T^2$ , thus completing the proof.

Let  $\tau$  be any sentence in  $T^2$ . We have to show that it is logically equivalent to a conjunction of sentences in  $S^2$ . Since  $S^1$  covers  $T^1$  and  $T^2$  is a subset of  $T^1$ , there are  $\sigma_1, \ldots, \sigma_k$  in  $S^1$  such that

$$\tau = \sigma_1 \wedge \cdots \wedge \sigma_k. \tag{2}$$

We know that  $T^2 \models S^2$ , since  $T^2$  covers  $S^2$ . We also know that  $T^2 \models \tau \models \sigma_i$ , for each  $\sigma_i$ . If some  $\sigma_i$  were not in  $S^2$ , the fact that  $S^2$  is a maximal subset of  $S^1$  not implying any sentence in  $\Sigma$  would entail that  $S^2 \cup \{\sigma_i\}$  implies some sentence  $\sigma \in \Sigma$ . But then  $T^2 \models \sigma$ , a contradiction. This shows that each  $\sigma_i$  is in  $S^2$ , and therefore  $S^2$  covers  $T^2$ .

Now let M be a model of some theory S' in the flock S'. By Condition 1, there is some theory T' in the flock T', such that S' covers T'. This implies that M is also a model of T'. Thus every model of S' is also a model of S'.

**Example 4:** The flocks  $\{\{A, B, A \land B\}\}$  and  $\{\{A, B\}\}$  are equivalent forever. The flocks  $\{\{A, B, A \lor B\}\}\}$  and  $\{\{A, B\}\}$  are not equivalent forever. If we delete A and then B, we get  $\{\{A \lor B\}\}\}$  from the first flock and  $\{\emptyset\}$  from the second one.

For arbitrary flocks we only have a sufficient condition for equivalence forever.

Theorem 8: Let S and T be two flocks that satisfy the conditions

$$(\forall S \in \mathbf{S})(\exists T \in \mathbf{T})(S \text{ covers } T \land T \text{ covers } S)$$
(3)

and

$$(\forall T \in \mathbf{T})(\exists S \in \mathbf{S})(T \text{ covers } S \land S \text{ covers } T)$$
(4)

Then S and T are equivalent forever.

**Proof:** See the proof of  $(6) \Rightarrow (2)$  in Theorem 7.

### Remarks:

- 1. By Theorem 6 we can replace "S covers T" in this theorem by the condition "for every subset of S, there is a logically equivalent subset of T."
- 2. The above conditions are not necessary for equivalence forever. For example, it is shown in [Ku] that the two flocks

$$S = \{\{A, B, A \equiv B\}, \{A, A \equiv B\}, \{B, A \equiv B\}\}$$

and

$$\mathbf{T} = \{\{A, A \equiv B\}, \{B, A \equiv L_i\}$$

are equivalent forever even though they do not satisfy Conditions 3 and 4...

## REFERENCES

- [BS] Bancilhon, F. and N. Spyratos, "Update Semantics of Relational Views," ACM Trans. on Database Systems 6 (1981), pp. 557–575.
- [CA] Carlson, C. R. and A. K. Arora, "The Updatability of Relational Views based on Functional Dependencies," Proc. IEEE COMPSAC, 1979, pp. 415-420.
- [DB] Dayal, U. and P. A. Bernstein, "Translation of Update Operations on Relational Views," ACM Trans. on Database Systems 8 (1982), pp. 381–416.
- [FUV] Fagin, R., Ullman, J. D., and M. Y. Vardi, "On the Semantics of Updates in Databases," Proc. 2nd ACM Symp. on the Principles of Database Systems, 1983, pp. 352-365.
- [J] Jacobs, B. E., "Application of Database Logic to the View Update Problem," Technical Report TR 960, Dept. of Computer Science, University of Maryland at College Park, 1980.
- [Ke] Keller, A. M., "Updates to Relational Databases through Views Involving Joins," In *Improving database usability and responsiveness* (P. Scheuermann, ed.), Academic Press, pp. 363-384.
- [Kl] Klug, A. C., "Theory of Database Mapping," Technical Report CSRG-98, Dept. of Computer Science, U. of Toronto, 1978.
- [Ko] Kowalski, R., "Logic as Database Language," Unpublished Manuscript, Dept. of Computing, Imperial College, London, 1981.
- [Ku] Kuper, G. M., "Semantics of Database Updates," Ph. D. Thesis, Stanford University, to appear.
- [NG] Nicolas, J. M. and Gallaire, H, "Database- theory vs. interpretation," In Logic and Databases (H. Gallaire and J. Minker, eds.), Plenum Press, 1978, pp. 33-54.
- [NY] Nicolas, J. M. and K. Yazdanian, "Integrity Checking in Deductive Databases," In Logic and Databases (H. Gallaire and J. Minker, eds.), Plenum Press, 1978, pp. 325-344.
- [O] Osman, I. M., "Updating Defined Relations," Nat'l Computer Conf., Vol. 48, AFIP Press, 1979, pp. 733-740.
- [R] Reiter, R., "Towards a Logical Reconstruction of Relational Database Theory," In On Conceptual Modelling: Perspectives from Artificial Intelligence, Databases, and Programming Languages (M. L. Brodie, J. Mylopoulos, and J. Schmidt, eds.), Springer-Verlag, 1984, pp. 191–233.
- [T] Todd, S., "Automatic Constraint Maintenance and Updating Defined Relations," Proc. IFIP 77 (B. Gilchrist, ed.), North-Holland, 1977, pp. 145-148.

î.